

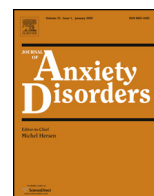
PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://repository.ubn.ru.nl/handle/2066/126858>

Please be advised that this information was generated on 2017-03-10 and may be subject to change.



Psychometric properties of reaction time based experimental paradigms measuring anxiety-related information-processing biases in children



H.M. Brown^a, T.C. Eley^a, S. Broeren^b, C. MacLeod^c, M. Rinck^d, J.A. Hadwin^e, K.J. Lester^{a,*}

^a Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College, London, United Kingdom

^b Department of Public Health, Erasmus Medical Centre, Rotterdam, Netherlands

^c Centre for the Advancement of Research on Emotion, University of Western Australia, Australia

^d Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

^e Department of Psychology, University of Southampton, United Kingdom

ARTICLE INFO

Article history:

Received 25 June 2013

Received in revised form 29 October 2013

Accepted 16 November 2013

Keywords:

Anxiety

Children

Information-processing

Psychometrics

Reliability

ABSTRACT

Theoretical frameworks highlight the importance of threat-related information-processing biases for understanding the emergence of anxiety in childhood. The psychometric properties of several tasks measuring these biases and their associations with anxiety were examined in an unselected sample of 9-year-old children ($N = 155$). In each task, threat bias was assessed using bias scores reflecting task performance on threat versus non-threat conditions. Reliability was assessed using split-half and test-retest correlations of mean reaction times (RTs), accuracy and bias indices. Convergence between measures was also examined. Mean RTs showed substantial split-half and test-retest correlations. Bias score reliability coefficients were near zero and non-significant, suggesting poor reliability in children of this age. Additionally, associations between bias scores and anxiety were weak and inconsistent and performance between tasks showed little convergence. Bias scores from RT based paradigms in the current study lacked adequate psychometric properties for measuring individual differences in anxiety-related information-processing in children.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY license.

1. Introduction

Cognitive models propose that anxiety is associated with a number of biases in information processing, including attentional biases for threatening information, the propensity to interpret ambiguous information as threatening and the tendency to avoid anxiety-provoking situations. The selective processing of threat and the related tendency to interpret ambiguity as threatening are argued to increase the likelihood of perceiving danger in the environment, where this process serves to cause or maintain anxiety (Muris & Field, 2008). Moreover, these information processing biases are suggested to lead to avoidant behaviour, which precludes opportunities to disconfirm threatening beliefs, thus

maintaining anxiety (Beck & Clark, 1997; Heuer, Rinck, & Becker, 2007).

1.1. Reaction time based information-processing paradigms

Most studies to date have measured information processing biases in anxiety using behavioural indicators that typically compare differences in reaction times (RTs) for emotional versus neutral stimuli to create bias scores (e.g., threat versus neutral words or pictures; Hadwin & Field, 2010). For example, a dot-probe task is used to measure selective attention or vigilance for threat in anxiety. This task requires participants to identify a probe (e.g., the location or number of dots) that follows one of two stimuli presented simultaneously (i.e., threat-neutral, positive-neutral or neutral-neutral stimulus pairs; MacLeod, Mathews, & Tata, 1986). Shorter RTs for probes following threat stimuli relative to probes replacing neutral stimuli in threat-neutral pairs indicate an attentional bias towards threat (threat vigilance). The reverse pattern of relative probe RTs indicates an attentional bias away from threat (threat avoidance). Several studies have found evidence to suggest anxiety-related selective attention to threat using this paradigm in children and adolescents (Roy et al., 2008; Telzer et al., 2008; Waters, Mogg,

* Corresponding author at: Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, 16 De'Crespigny Park, London SE5 8AF, United Kingdom. Tel.: +44 020 7848 5414; fax: +44 020 7848 0866.

E-mail address: kathryn.lester@kcl.ac.uk (K.J. Lester).

Bradley, & Pine, 2008), although others find that childhood anxiety is associated with avoidance of threat (Brown et al., 2013; Stirling, Eley, & Clark, 2006) and yet others find no associations with anxiety (Waters, Lipp, & Spence, 2004).

Selective attention to threat can also be assessed using tasks designed to measure inhibitory control, such as emotional variants of Stroop and Garner tasks (Gilboa-Schechtman, Ben-Artzi, Jeczemien, Maro, & Hermesh, 2004). In emotional Stroop tasks, participants are asked to identify the colour of a stimulus (e.g., a word or picture outline), while ignoring its emotional meaning. Similarly, the Garner task requires individuals to identify certain non-emotional stimuli properties (i.e., gender) whilst ignoring emotionally valenced stimulus properties (e.g., facial expression). In both versions of this paradigm, an attentional bias for threat is inferred when latencies to process non-emotional stimulus dimensions are longer for threatening than neutral stimuli. Some studies assessing inhibitory control have found increased interference of angry faces on colour matching in children with elevated anxiety (Hadwin, Donnelly, Richards, French, & Patel, 2009; Klein, Becker, & Rinck, 2011; Morren, Kindt, van den Hout, & van Kasteren, 2003; Richards, French, Nash, Hadwin, & Donnelly, 2007).

Other tasks have explored attention processes linked to threat detection. Visual search tasks, for example, require participants to search for threat and non-threat stimuli typically within an array of neutral distracters. RTs to find target stimuli are used as a measure of detection or hypervigilance for threat (Donnelly, Hadwin, Menneer, & Richards, 2010). Links between threat detection and anxiety are typically expressed as a negative association between anxiety and a slope gradient that reflects changes in RT as the number of distractor stimuli in a search increases (Hadwin et al., 2003). Studies with children and adolescents have found that young people are faster to detect threat (versus non-threat stimuli) as depicted in angry faces (Perez-Olivas, Stevenson, & Hadwin, 2008) and show increased efficiency when making decisions about the absence of threat (Hadwin et al., 2003).

Morphed face tasks were developed to consider anxiety-linked differences in RTs and errors in deciding when dynamic or static faces display positive or negative emotions (Joormann & Gotlib, 2006). In static morph paradigms, participants are presented with faces of emotional expression (e.g., anger, fear, sadness, disgust, happiness; Lau et al., 2009) varying in intensity, whilst in dynamic tasks, short videos of faces gradually transform (morph) from neutral to prototypical emotional expressions. Studies with children have shown that elevated anxiety is associated with increased misattributions of anger to faces with low levels of emotional information in a static morphed faces in 10 year-olds (Richards et al., 2007). However, other studies using dynamic morph tasks have found no associations between self-report anxiety and RTs or accuracy on a morph task in similarly aged children (Lau et al., 2009), and further studies have identified anxiety-related effects only in older and not younger children when using latent class regression on data from 4 to 12 year olds (Broeren, Muris, Bouwmeester, Field, & Voerman, 2011).

Most RT tasks have typically focused on selective attention or detection of threat stimuli in child and adolescent anxiety. However, more recent paradigms (e.g., the approach–avoidance task or AAT) have considered anxiety-related behavioural approach and avoidance responses to positive and negative emotional stimuli (Chen & Bargh, 1999). One technique involves measuring the relative speed of pull (approach) and push (avoidance) arm movement responses (using a computer joystick) to emotional and neutral stimuli. Approach and avoidant behaviours are inferred from participants' relative speed to execute push and pull responses to different stimuli (Marsh, Ambady, & Kleck, 2005). Preliminary studies using this task have demonstrated behavioural avoidance of

spider pictures in adult spider phobics (Rinck & Becker, 2007) and behavioural avoidance of emotional (angry and happy) faces in socially anxious adults (Heuer et al., 2007), as evidenced by faster RTs for pushing than pulling these stimuli. Similarly, girls but not boys (aged 9–12 years) with high self-reported spider fear showed faster RTs for pushing than pulling of spider pictures using an AAT, indicating behavioural avoidance (Klein, Becker, & Rinck, 2011a, 2011b).

As highlighted in previous sections, the findings from information processing studies with child samples have proven somewhat inconsistent. The mixed pattern of findings within and across information processing tasks may in part be explained by increased variability in the methodologies adopted to assess information processing biases in children compared to adults. For example, child research has often been marked by greater variability in terms of sample characteristics and anxiety informant (e.g., self-, parent- or clinician-ratings) as well as differences in task format and stimuli, as researchers tweak task parameters to address their research questions while meeting the demands associated with testing younger participants (Field & Lester, 2010).

Age-related effects may also impact on the development and measurement of information processing biases. It is possible that cognitive developmental factors may influence the emergence of information processing biases during childhood (Field & Lester, 2010). Alternatively, developmental changes in cognitive processes may mediate performance on experimental tasks indexing these biases rather than the biases themselves. A number of changes are noted in cognitive processes across development, including advances in attentional and inhibitory control and emotional recognition (Klenberg, Korkman, & Lahti-Nuuttila, 2001). However, the paucity of research explicitly examining developmental trends in anxiety-related information processing and the mixed results across studies focusing on different age groups has meant that developmental effects on information processing are not well understood. As a result, findings from studies with adults cannot simply be extended to children and studies explicitly examining the extent to which bias indices from these tasks are reliably stable and valid in children are required.

1.2. Reliability and temporal stability

The focus on information-processing biases as possible factors that cause or increase anxiety has led to increased use of these paradigms as indicators of treatment outcome in anxiety (Mathews, Mogg, Kentish, & Eysenck, 1995; Mattia, Heimberg, & Hope, 1993) and more recently as possible treatment methods themselves (e.g., attentional bias modification; Hakamata et al., 2010). Researchers and clinicians therefore need to be confident that cognitive paradigms are reliably stable over time, so that differences in task performance can be attributed to a change in cognition and not random fluctuations in measurement over time. However, studies are yet to consider the reliability of RT based paradigms developed to measure threat-related information processing in child and adolescent populations.

1.3. Convergence between paradigms

In addition, theoretical frameworks in anxiety suggest that biases in information-processes measured using different paradigms should be linked and that those measuring early processes should be associated with those that reflect later processing (Daleiden & Vasey, 1997; In-Albon & Schneider, 2010; Muris & Field, 2008). Some studies have found convergence in performance between different attentional tasks. For example, Richards et al. (2007) showed that a lack of inhibitory control to threat in an emotional Stroop task was linked to difficulties in

emotion discrimination in a morphed face task in late childhood. However, further studies have shown no association between inhibitory control in an emotional Stroop task and attentional vigilance in a dot probe task (Dalglish et al., 2003) or between vigilance for threat in a dot-probe tasks and emotion discrimination in a morph task (Broeren et al., 2011).

1.4. The current study

Further research is needed to consider stability in performance on information-processing tasks over time and convergence across a broader range of tasks. The current study therefore examined the psychometric properties of a range of information-processing tasks and their associations with anxiety in a large unselected sample of children aged 8–10 years. This relatively narrow age range was selected as middle childhood is argued to represent a key period in the emergence of information-processing biases (Field & Lester, 2010) and is prior to the mean age of onset of anxiety disorders (Kessler et al., 2005). A relatively tight age focus also went some way to circumventing possible age-related effects on information processing, which were not a focus of the present study. Furthermore, the tasks selected to assess information processing biases in the current study were appropriate for 8–10 year olds but would not have been suitable for younger children. The psychometric properties of prototypical variants of a selection of widely used measures of selective attention (dot-probe), detection (visual search), inhibitory control (emotional Stroop and Garner tasks) and emotion discrimination (emotional morph task) were assessed. Novel task variants were also included to measure selective attention (missile-probe) and behavioural avoidance (AAT). In all tasks, threat was depicted using angry faces and behavioural responses (RTs) to these stimuli were compared to baseline (neutral faces) or positive (happy faces) conditions to create bias scores for the processing of threat information. Emotional faces were chosen to reflect the emergence of social themes of threat in children in middle childhood in the normal developmental trajectory of fear (Gullone, 2000). In line with theoretical frameworks in anxiety, we explored whether mean condition RTs and bias scores on information-processing tasks demonstrated stability within each testing session and over time using split-half and test-retest reliability estimates, respectively. Associations with anxiety were examined by correlating information-processing bias scores with children's self-report anxiety at each testing session. Following previous research, it was anticipated that attention processes in each task should be linked, especially for tasks designed to measure the same underlying process as with emotional Stroop and Garner tasks. We examined this by correlating bias scores from each task with one another.

2. Methods

2.1. Participants

Ethical approval was granted by the Psychiatry, Nursing and Midwifery Research Ethics Subcommittee of King's College London (ref no: PNM/10/11-37). Two primary schools were recruited based on a number of Ofsted criteria (e.g., fewer than average students with English as an additional language or with learning difficulties) to ensure the sample represented socioeconomic distributions in the general population. Parents of children aged 8–10 years were sent an information sheet, brief family background questionnaire and consent form. Children of consenting parents were introduced to the study and gave verbal assent.

The initial sample consisted of 155 children (67 males, 88 female); 32 from the first school and 123 from the second; 36%

and 68% response rates, respectively (Fig. 1). Due to other pressures within the school, the first school had a lower response rate and ceased their participation after completion of the first wave of data collection. Of participating children, 78% were classified as Caucasian, slightly less than in the general population (93%; Scott, Pearce, & Goldblatt, 2001) and all spoke English as their first language.

Following wave 1, parents of participating children were invited to re-consent for two additional testing waves, resulting in retention of 107 children from the original sample at wave 2 and 104 at wave 3. Children who dropped out were slightly older than those retained; $t(121) = 2.49$, $p < .01$ but did not differ in sex, ethnicity or anxiety level. Reasons for withdrawal included concern regarding time missed from lessons or significant changes in their child's personal circumstances.

2.2. Measures

2.2.1. Revised Children's Anxiety and Depression Scales (25-item version)

Revised Children's Anxiety and Depression Scales (25-item version) (RCADS-25; Muris, Meesters, & Schouten, 2002) comprises 25-items measuring common symptoms of anxiety and depression. Children rated how often (never, sometimes, often, always) they experienced each item. Only anxiety items are used in the current analyses. Responses were coded 1–4 and summed across anxiety items to create total scores. Higher scores indicate greater anxiety symptom severity. Internal consistency ($\alpha s = .87-.95$) and test-retest reliability coefficients ($r s = .78-.86$, $p s < .001$) were substantial at all time points.

2.2.2. Information-processing bias paradigms

Dot-probe task. Thirty-two models portraying angry, happy and neutral facial expressions were selected from the NimStim face set (Tottenham et al., 2009). Equal numbers of males and females were used. Face pairs comprised two pictures of the same model presented horizontally next to one another representing angry-neutral and happy-neutral pairings. Two test blocks were administered, each comprising 96 randomly presented trials. Blocks were separated by a self-determined break. Each dot-probe trial consisted of a centrally-positioned fixation cross presented for 1000–2000 ms which was replaced by a face-pair presented for 500 ms. A probe then appeared consisting of either one or two dots in a location corresponding to the centre of one of the previously presented faces. Probe type was counterbalanced. Equal numbers of each probe appeared in the location of each face type. Participants indicated as quickly and accurately as possible how many dots were displayed by pressing corresponding response box buttons. Probes remained on screen until participants responded. RTs and accuracy of responses were recorded. Bias scores were calculated by subtracting mean RTs for probes presented in the locus of the emotional image from mean RTs for probes presented in the locus of the neutral image for both angry-neutral and happy-neutral trials. Positive bias scores indicate a bias towards preferentially processing emotional stimuli whilst negative scores indicate avoidance of such stimuli.

Missile-probe task. The missile-probe paradigm (MPT) is a novel adaptation of the dot-probe task. In this variant, probe exposure duration was calibrated 'on-line' to maintain an average response accuracy of around 75%, enabling analysis of differential error rates across conditions rather than relying entirely on response latencies, which show greater variability in children and thus may represent a source of unreliability in traditional dot-probe data (Broeren et al., 2011). Additionally, children were awarded points for each correct response to maintain participant motivation, in order to reduce response variability due to waning engagement with the task.

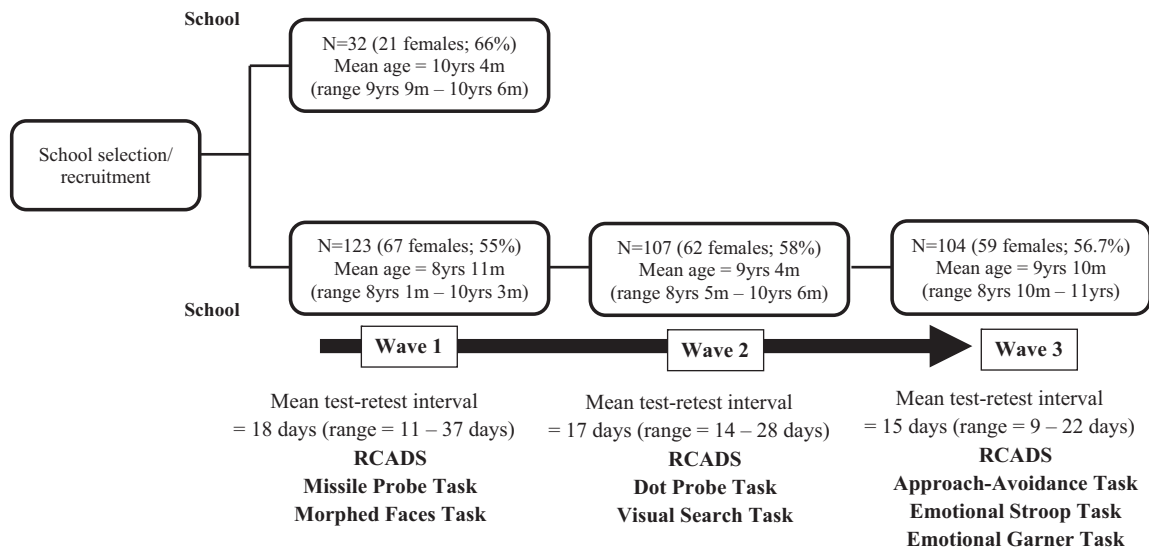


Fig. 1. Schematic diagram of study sample.

Thirty-two models portraying angry, happy and neutral expressions were selected from the Radboud faces database (Langner et al., 2012). Equal numbers of male and female models were used. Sixty-four emotion-neutral face pairs were created with equal angry-neutral and happy-neutral pairings. The MPT consisted of one practice block and two test blocks; each with 64 trials. The practice block contained neutral-neutral face pairings only. In each test block emotion-neutral face pairs were selected so that each set of 16 trials contained one of each possible unique trial combination (emotional face [angry or happy], emotional face location [left or right], probe location [left or right] and probe direction [left or right]). Trials began with the child's score presented in the centre of the screen for 1000 ms serving as a fixation point followed by a face-pair presented for 1000 ms (Fig. 2). A 'missile' probe, pointing either left or right, then appeared in the position corresponding to the centre of one of the faces. Next the probed and unprobed spaces were occluded by a picture of a cloud. Children indicated which direction the missile was pointing (left or right) by pressing corresponding laptop keys. Incorrect responses were followed by a sad trombone sound. Correct responses were followed by the missile visually exploding with an accompanying explosion noise. Children were awarded 15 points for correct responses made within one second of the probe display onset, 10 points for correct responses between one and two seconds and five points for correct responses over two seconds. An average response accuracy of approximately 75% was maintained by altering the duration for which the probe was exposed before being occluded by clouds. In order to determine the initial probe duration for test blocks, probe durations in the practice trials began at 1000 ms and got 100 ms shorter or 40 ms longer in response to correct and incorrect responses, respectively. The starting duration for test trials was determined as the average exposure duration across the final 16 practice trials. Throughout test blocks

response accuracy was monitored and probe duration was recalibrated every 16 unique trial combinations. Probe duration reduced by 20 ms if accuracy was greater than 13, and decreased by 20 ms if less than 10 responses were correct, across these 16 trials. The lower limit of exposure duration was 10 ms. RTs and accuracy of responses were recorded. Bias scores were calculated for RTs in the same way as for the dot-probe task so that positive bias scores indicated attentional vigilance for emotional stimuli and negative scores indicated avoidance. Similarly, accuracy bias scores were calculated by subtracting mean accuracy for probes appearing behind angry faces from mean accuracy for probes appearing in place of neutral faces. Positive accuracy bias scores indicate greater accuracy for probed angry relative to probed neutral faces.

Visual search. The visual search task was adapted from that used by Hadwin et al. (2003), consisting of three test blocks comprising 72 trials each representing searches for angry, happy and neutral schematic faces. In half the trials the target face was present and in half it was absent. In target present trials, the face was presented amongst an array of distracters consisting of non-facial reconfigurations of the constituent features of the target face (i.e., scrambled face). Target absent trials contained only distracters. Stimuli arrays contained 4, 6 or 8 stimuli arranged equidistantly in a circle. Each trial was preceded by a centrally-positioned fixation cross presented for 1000–2000 ms. Children indicated as quickly and accurately as possible whether the target face was present or absent by pressing corresponding response box buttons. Stimuli remained on screen until responses were made. RT and accuracy of response were recorded for each trial. Search slopes and intercepts across increasing set size were calculated for RTs for each search condition. The gradient of the slope indicated the extent to which RTs increased across increasing set size. The intercept showed the point at which the slope crossed the Y axis, indicating overall

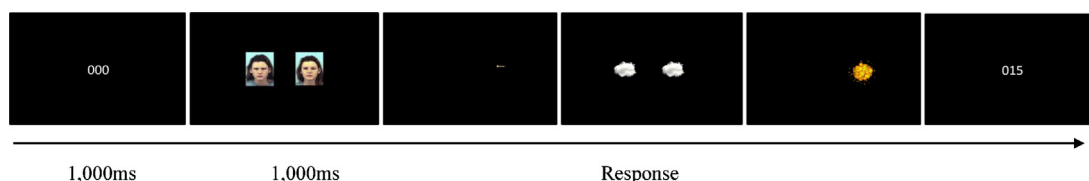


Fig. 2. Schematic representation of the missile-probe task.

difficulty on each search condition (Hadwin et al., 2003). Bias scores were created by subtracting search slopes for neutral faces from both angry and happy search slopes for both target present and target absent trials. Positive scores indicate a greater search 'cost' across increasing set size for emotion relative to neutral searches, whilst negative scores indicate less relative 'cost' of increasing set size on searches for emotional target searches.

Emotional Stroop task. Eight models comprising equal numbers of male/female and adult/child faces displaying angry, happy and neutral emotions were selected from the Radboud database (Langner et al., 2012). Green, yellow, blue and red tinted versions of each picture were created; a total of 96 trials. The task began with a practice block of eight neutral trials. Adult and child images were divided into separate test blocks with a self-determined break in between. Test block order was randomly determined. Trials were presented in a pseudo-random order so that no two images of the same colour were seen consecutively. Each trial began with a centrally-positioned fixation cross presented for 1000–2000 ms followed by a face image. Children indicated as quickly and accurately as possible the colour of the face by pressing corresponding response box buttons. Faces remained on screen until responses were made. RTs and accuracy of response were recorded. Bias scores were calculated by subtracting RTs for neutral stimuli from RTs for emotional (angry/happy) stimuli. Positive scores indicate greater interference of emotional relative to neutral stimuli.

Garner task. Similar to the emotional Stroop, eight unique models comprising equal numbers of male/female and adult/child faces displaying angry, happy and neutral emotions were selected from the Radboud database (Langner et al., 2012). Each image was presented three times; a total of 72 trials. The task began with a practice block of eight neutral trials. Adult and child images were divided into separated test blocks separated by a self-determined break. Test block order was randomly determined. Trials were presented in a pseudo-random order so that no two trials of the same model were seen consecutively. Each trial began with a centrally-positioned fixation cross presented for 1000–2000 ms followed by a face image. Children indicated as quickly and accurately as possible the gender of the face by pressing corresponding response box buttons. Faces remained on screen until responses were made. Response latencies and accuracy were recorded. Bias scores were calculated by subtracting RTs for neutral stimuli from RTs for emotional (angry/happy) stimuli. Positive scores indicate greater interference of emotional relative to neutral stimuli.

The face morphing task (Broeren et al., 2011). Twenty models comprising equal numbers of males and females and open and closed mouth facial expressions were selected from the NimStim database (Tottenham et al., 2009). Each model's neutral expression was morphed ("MorphMan 4.0," 2003) in 75 increments of increasing emotional intensity with both their angry and happy expressions (Fig. 3), creating a total of 40 dynamic morphs. Each unique morph was presented once, resulting in 40 trials. Each trial lasted 10 s. Trials were separated by a fixation cross displayed for 1000–2000 ms. Children indicated which emotional expression was being displayed by the face, by pressing the corresponding response box buttons as soon as the identity of this emotional expression became evident to them. Upon this response the video stopped and the next trial began. RTs to make a response and accuracy of response were recorded. Bias scores were calculated by subtracting mean RTs for angry trials from mean RTs for happy trials. Positive scores indicate speeded detection of angry relative to happy facial expressions.

Approach–avoidance task (AAT). Sixteen models displaying angry, happy and neutral facial expressions were chosen from the Radboud database (Langner et al., 2012). Equal numbers of male and female and of child and adult models were used. Sepia and greyscale versions of each image were created in seven different sizes

(76×91 , 106×130 , 154×185 , 220×263 , 314×377 , 449×535 , 642×768 pixels); a total of 96 trials. Each trial began with a face image (220×263 pixels) presented centrally on the computer screen. Participants pushed the joystick (Logitech Attack 3) for grey faces and pulled for sepia faces. Image size decreased and increased for push and pull movements respectively, giving the impression of the face moving further away or closer. Images remained on screen until the joystick was moved fully in the correct direction. Participants began with 10 practice trials of neutral faces and then two test blocks of 96 trials; 12 trials per emotion condition, per block (e.g., 12 sepia happy trials, 12 sepia angry trials, etc.) and 12 neutral filler trials (6 per colour shading). Trial order was pseudo-randomised and fixed across participants. RTs to make an initial response and accuracy of the initial responses were recorded. Bias scores were created by subtracting mean RTs for 'compatible' conditions (e.g., pulling positive expressions, pushing negative expressions) from the respective 'incompatible' conditions (pushing positive expressions, pulling negative expressions). Bias scores were also calculated on neutral trials by subtracting RTs for "pull" trials from "push" trials as an indication of individual tendencies to push and pull faces.

2.3. Procedure

Questionnaires and information-processing paradigms were programmed in E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA), apart from the MPT which was programmed in BBC Basic, and the AAT which was programmed in Microsoft Visual Basic.

The study consisted of three data collection waves (Fig. 1). Each wave consisted of two testing sessions approximately 2–3 weeks apart; mean test–retest intervals within each wave were 18 (range = 11–37 days), 17 (14–28) and 15 (9–22) days for waves 1, 2 and 3, respectively. Intervals between waves varied in line with the school's academic calendar. RCADS-25 and a range of information-processing paradigms were completed at each wave. RCADS-25 was always completed first. The order of experimental tasks was counterbalanced across participants using a Latin square design. Individual's task order was identical for both sessions within a wave. Testing sessions lasted no more than one hour. Children were seen individually in a quiet classroom and were supervised by a researcher throughout data collection. Instructions and questionnaire items were read aloud to ensure comprehension. Children received a craft gift at the end of waves 1 and 2 and a book voucher at the end of wave 3.

2.4. Analyses

Mean RTs were calculated for each condition on each task, removing incorrect responses and data values above or below 2.5 standard deviations from individual means and <100 ms.¹ Participants with more than 25% errors or outliers on a single task condition were excluded for that task, apart from on the MPT where the task was designed to keep accuracy around 75%. All variables were found to approximate normal distribution so parametric analyses were used throughout. Reliability was assessed for RTs from dot-probe, missile-probe, Stroop, Garner, Morph and AAT tasks; slope variables from the visual search task and accuracy data from the missile-probe task, as well as their respective bias scores.

¹ One trial from the VS and AAT tasks was excluded from calculations of trial means owing to a programming error on each of these tasks. Specifically, a wrong trial image was programmed in the angry child trials of the AAT and one of the distracter stimuli in the VS angry present trials was of a different formation to other distracters owing to an editing error.



Fig. 3. Schematic of the face morphing task.

Internal reliability was measured by computing within-subjects split-half correlations with Spearman–Brown corrections for summary scores from each task. Test–retest reliability was assessed by correlating summary scores from each task at sessions with respective scores at session two. To examine associations with anxiety, bias scores from each information-processing paradigm at each testing session were correlated with anxiety scores from the same session. Convergence between measures was assessed by correlating bias scores from each of the bias paradigms with each other at both testing sessions. Bonferroni corrections were applied to account for multiple comparisons.

3. Results

3.1. Anxiety

Mean anxiety scores and their variances were comparable to those reported in previous unselected samples (Table 1). The proportion of children with clinically elevated anxiety was identified using the normative cut-off values for the top 25% and 10% of total anxiety scores recommended for the RCADS-25; scores of over 16 and 26, respectively (Muris et al., 2002). At the first time of measurement, 41% and 13% of children, respectively, exceeded these scores, indicating frequencies of clinically elevated anxiety in line with expectations for unselected samples. Anxiety decreased significantly across measurement; $F(5,510) = 17.97$, $p < .001$. Internal consistency and test–retest reliability were substantial across all time points (moderate .3–.5, substantial >.5; Field, 2005).

3.2. Psychometric properties of information-processing paradigms

With the exception of the MPT, error and outlier rates were uniformly low across all tasks as expected and resulted in a total deletion of <1% of all data points. As a result, we were unable to examine the psychometric properties of accuracy data on these tasks. Table 2 reports descriptive statistics and reliability estimates for summary scores from each task at each testing session. RTs and bias scores from emotional Stroop, Garner and AAT tasks were collapsed across adult and child stimuli as ANOVAs revealed no significant response differences. There were significant main effects of testing session on RTs from morph; $F(1,140) = 35.49$, $p < .01$; $s_1 = 4416.03$ (SD = 102.06) ms, $s_2 = 3909.37$

(84.79) ms, visual search; $F(1,102) = 115.79$, $p < .01$; $s_1 = 1897.97$; $s_2 = 1614.01$ (40.93) ms and MPT tasks; $F(1,143) = 22.31$, $p < .01$; $s_1 = 1002.20$ (25.28) ms; $s_2 = 895.01$ (22.48) ms. RTs were faster at the second measurement, indicating possible practice effects. There were no significant differences for dot-probe, emotional Stroop, Garner or AAT tasks or MPT accuracy scores. Large standard deviations were observed across all tasks (2nd and 3rd columns in Table 2), indicating substantial response variability across participants. These were comparable to variability seen in previous studies with children but considerably larger than those seen in adult studies (e.g., standard deviations for trial RTs on a dot-probe tasks ranging from 62 to 91 ms in adults compared to 137–160 ms in children; Waters et al., 2004).

3.2.1. Reliability

Split-half correlations with Spearman–Brown corrections demonstrated substantial internal consistency for mean RTs across dot-probe, missile-probe, emotional Stroop, Garner, morph and AAT tasks ($r_s = .63$ –.91) and generally for accuracy scores on the MPT ($r_s = .11$ –.42) whilst split-half estimates were somewhat lower but still significant for visual search slope variables ($r_s = .25$ –.54). Conversely, split-half correlations for bias scores were largely non-significant and unacceptably low ($r_s = -.24$ to .33). Exceptions included bias scores from angry–neutral slopes on the visual search task ($r_s = .22$ and .38 for times 1 and 2, respectively) and angry–happy bias scores on the morph task ($r_s = .30$ and .51) which showed moderate internal consistency. Results were similar when examining test–retest reliability. Mean RTs indicated moderate to substantial reliability across sessions ($r_s = .43$ –.75 across all tasks) although test–retest correlations for visual search slopes were somewhat lower but still significant ($r_s = .24$ –.36). Conversely, test–retest reliability coefficients for bias scores were largely non-significant and near zero ($r_s = -.06$ to .33) with the exception of the morph bias index ($r = .33$) and angry–neutral slope bias on target present trials on the visual search ($r = .21$).

3.2.2. Associations with anxiety

Associations between bias scores and anxiety and between bias scores from different tasks are shown in Table 3. Correlations between the first measurements are shown below the diagonal. Correlations between the second measurements are shown above. After applying a Bonferroni correction ($\alpha = .05/32 = .0016$), no associations between bias scores and anxiety remained significant.

Table 1
Descriptives of anxiety scores.

	N	Range	Mean (SD)	Internal consistency (α)	Test–retest (r)
<i>RCADS^a anxiety</i>					
w1s1	155	1–51	18.71 (10.45)	.89	.86*
w1s2	155	0–57	17.61 (11.02)	.91	
w2s1	107	1–39	16.29 (8.87)	.87	.78*
w2s2	106	0–34	15.15 (9.25)	.88	
w3s1	103	0–37	13.66 (8.30)	.87	.78*
w3s2	104	0–34	13.14 (8.47)	.89	

^a RCADS-25 RCADS Revised Children's Anxiety and Depression Scales w1/2/3 wave 1/2/3; s1/2 session 1/2.

* $p < .001$.

When examining the magnitude of effect ignoring the relatively stringent alpha level, associations rarely exceeded .2 indicating that bias indices were at best weakly associated with anxiety. For example, there was a weak, consistent association between the morph bias index and anxiety, indicating anxiety was associated with slower identification of angry relative to happy faces. However, this was not significant at the .05 level. Additionally, associations were not consistent across testing sessions with some indicating opposite directions of effect; for example, associations with happy–neutral bias scores on the emotional Stroop task ($r = -.25$ and .18 at sessions 1 and 2, respectively).

3.2.3. Convergence between measures

There were no significant associations between bias scores across different information-processing tasks after correcting for multiple comparisons ($\alpha = .05/240 = .0002$). However, different bias indices taken from the same measure were moderately associated even after stringent corrections for multiple comparisons. Specifically, visual search bias scores were all positively associated with each other at session one but were more inconsistent at session two and correlations were moderate at best suggesting limited convergence. Of particular interest given previous studies, bias scores from dot-probe and emotional Stroop tasks ($r_s = -.13$ to .12), from missile- and dot-probe tasks ($r_s = -.13$ to .10) and from emotional Stroop and Garner tasks ($r_s = -.04$ to .13) were largely uncorrelated, suggesting no convergence between these tasks despite putatively measuring the same construct.

4. Discussion

The aim of the current study was to examine the psychometric properties of a selection of widely used measures of anxiety-related information processing biases in a large unselected sample of children. The results showed that bias scores from a range of information-processing tasks demonstrated poor internal and test–retest reliability, were not strongly or consistently associated with anxiety and showed little convergence with one another.

4.1. Reliability

Split-half and test–retest correlations of mean condition RTs for all tasks, accuracy scores from the MPT and search slopes from the visual search task demonstrated moderate to substantial consistency both within and across testing sessions. Whilst the magnitude of reliability coefficients for mean condition RTs, accuracy (MPT) and slope (visual search) scores suggest adequate reliability, they do not capture differentials in processing between trials presenting stimuli of opposing emotional valence and so do not index information-processing biases. Instead, high correlations for condition summary scores within and across sessions could reflect consistency in general processing tendencies; that is, some people are systematically faster or slower than others, regardless of emotional condition.

In contrast to estimates for condition summary scores, reliability coefficients for bias scores from these tasks were generally small. Emotional Stroop, Garner, AAT and the majority of dot-probe and missile-probe bias scores yielded near-zero reliability coefficients, indicating poor reliability. Some visual search (e.g., angry–neutral absent slopes) and dot-probe (happy–neutral RT) bias scores showed evidence of internal consistency but not test–retest reliability, whilst the morph bias index demonstrated moderate internal and test–retest reliability. However, even for these tasks reliability estimates were near the accepted lower limit (.5; Field, 2005) and varied across testing sessions.

These results suggest that either the current tasks are not reliable when used in middle childhood or the processes they were designed to measure are not temporally stable (i.e., are not trait-like). This has possible implications for theoretical frameworks in anxiety which suggest anxiety-related information processing biases represent stable trait-like characteristics, which play a role in the maintenance of anxiety (Muris & Field, 2008). However, the current study suggests that information-processing biases, at least when measured using RT based tasks, are not reliably stable over a 2–3 week period in children aged 8–10 years. Temporal instability of these tasks also has implications for recent research which uses paradigms similar to those in the current study to measure treatment outcomes in anxiety (Mathews et al., 1995; Mattia et al., 1993). Low temporal stability means that researchers and clinicians cannot be sure that differences in task performance can be attributed to change in cognition and not stochastic fluctuations in measurement over time.

However, it is worth noting that reliability estimates for bias scores will always be lower than those for mean RTs from the constituent conditions (e.g., angry and neutral trials in angry–neutral bias scores). This is because measurement error from the two trials is compounded when combined into a single index, resulting in attenuation of correlation coefficients (Overall & Woodward, 1975). Measurement error may be particularly significant when using RT based tasks with children where RTs are more variable than when using similar tasks with adults (Waters et al., 2004). However, studies examining the psychometric properties from various task variants when used with adults have also revealed low test–retest correlations for bias scores (Eide, Kemp, Silberstein, Nathan, & Stough, 2002; dot-probe; Schmukle, 2005; Staugaard, 2009; emotional Stroop; Strauss, Allen, Jorgensen, & Cramer, 2005) suggesting possible methodological problems unrelated to age.

4.2. Associations with anxiety

In general, bias scores in the current study were not strongly or consistently associated with self-reported anxiety. The general lack of associations with anxiety suggests that either information-processing biases are not associated with self-report anxiety in unselected children aged 8–10 years or that current tasks are insufficiently sensitive to detect such effects. The former possibility has received some support. A meta-analysis reported that only

Table 2
Descriptives and reliability coefficients for information-processing paradigms.

	Mean (SD) (ms)		Split-half correlation ^a (r)		Test-retest correlation ^b (r)
	T1	T2	T1	T2	
<i>Dot-probe (N = 103/105)</i>					
Angry RT	775.59 (133.37)	762.55 (160.74)	.87**	.85**	.62**
Neutral RT	807.17 (140.55)	778.27 (158.84)	.91**	.87**	.67**
Angry-neutral bias	31.58 (67.31)	15.72 (67.54)	.29*	.12	-.06
Happy RT	782.40 (139.05)	764.84 (166.38)	.84**	.88**	.66**
Neutral RT	788.95 (145.33)	772.71 (159.14)	.88**	.87**	.64**
Happy-neutral bias	6.55 (57.19)	7.87 (76.66)	.16*	.33*	-.02
<i>Missile-probe (N = 150/149)</i>					
Angry RT	991.89 (299.82)	898.55 (306.56)	.84**	.79**	.52**
Neutral RT	1010.11 (316.80)	892.39 (277.98)	.82**	.65**	.50**
Angry-neutral bias	18.21 (121.34)	-6.16 (166.80)	.27**	.19*	.02
Happy RT	1001.11 (316.30)	892.58 (277.34)	.82**	.80**	.52**
Neutral RT	1002.59 (323.05)	898.79 (277.86)	.85**	.79**	.50**
Happy-neutral bias	1.47 (156.68)	6.22 (152.36)	.09	.10	.14
Angry accuracy	.78 (.10)	.77 (.09)	.35**	.17*	-.01
Neutral accuracy	.78 (.11)	.76 (.09)	-.26**	.11	.14
Angry-neutral bias (accuracy)	-.01 (.09)	-.01 (.10)	-.11	-.16	.11
Happy accuracy	.81 (.10)	.79 (.09)	.30*	.21*	-.07
Neutral accuracy	.78 (.11)	.75 (.10)	.42*	.25**	.04
Happy-neutral bias (accuracy)	-.02 (.11)	-.04(.10)	.02	.04	.16
<i>Visual search (N = 97/96)</i>					
Happy absent slope	185.19 (114.43)	145.24 (96.23)	.40**	.54**	.24**
Angry absent slope	216.01 (119.44)	173.75 (103.23)	.41**	.52**	.36**
Neutral absent slope	117.08 (131.07)	126.98 (94.19)	.25**	.46**	.36**
Angry-neutral absent bias	45.04 (157.43)	46.77 (118.58)	.38**	.22**	.15
Happy-neutral absent bias	14.12 (150.37)	18.25 (100.27)	.12	.10	.02
Happy present slope	92.87 (86.29)	58.11 (70.50)	.15	.69**	.06
Angry present slope	74.29 (72.91)	55.40 (95.67)	.16	.47**	.21*
Neutral present slope	71.54 (67.96)	51.76 (66.72)	-.06	.15	.02
Angry-neutral present bias	2.74 (108.06)	3.64 (128.62)	.12	.32**	.21*
Happy-neutral present bias	21.33 (101.06)	6.34 (82.13)	.07	.06	-.09
<i>Emotional Stroop (N = 103/102)</i>					
Angry RT	986.78 (183.06)	964.02 (189.90)	.78**	.85**	.65**
Happy RT	986.22 (194.98)	966.63 (191.87)	.87**	.89**	.74**
Neutral RT	980.47 (182.25)	957.95 (188.38)	.81**	.83**	.73**
Angry-neutral bias	6.31 (96.66)	6.06 (89.93)	.07	.13	.13
Happy-neutral bias	5.75 (82.29)	2.68 (78.54)	.05	.01	.16
<i>Garner (N = 102/103)</i>					
Angry RT	793.49 (178.52)	810.00 (266.90)	.74**	.82**	.62**
Happy R	801.07 (187.18)	812.08 (211.05)	.75**	.63**	.67**
Neutral RT	812.11 (183.18)	810.77 (255.53)	.78**	.71**	.68**
Angry-neutral bias	18.65 (96.81)	.78 (91.56)	-.04	-.24*	-.07
Happy-neutral bias	11.03 (82.30)	-1.31 (126.35)	-.11	-.15*	.25*
<i>Morph (N = 134/139)</i>					
Angry RT	4685.57 (1097.62)	4103.01 (9751.55)	.94*	.92*	.65*
Happy RT	4384.49 (1167.09)	3845.06 (1056.85)	.95*	.92*	.70*
Angry-happy bias	-301.07 (549.71)	-257.95 (648.80)	.30*	.51*	.33*
<i>AAT (N = 102/103)</i>					
Angry: push RT	973.86 (186.76)	887.64 (213.58)	.68**	.78**	.75**
Angry: pull RT	958.10 (209.09)	894.95 (272.26)	.77**	.76**	.79**
Angry bias	15.76 (133.92)	-7.31 (134.66)	.14	.08	.13
Happy: push RT	940.74 (182.79)	879.19 (254.57)	.61**	.84**	.73**
Happy: pull RT	969.27 (193.07)	874.24 (230.86)	.67**	.88**	.78**
Happy bias	-28.53 (114.28)	4.95 (97.82)	-.03	-.36**	.06
Neutral: push RT	973.35 (188.83)	878.67 (220.92)	.67**	.88**	.83**
Neutral: pull RT	973.90 (201.38)	881.57 (224.39)	.57**	.85**	.67**
Neutral bias	-.55 (120.09)	-2.90 (108.75)	-.01	.34**	.06

* $p < .05$; ** $p < .01$; ms: milliseconds, (r) Pearson's correlation co-efficient, AAT: approach-avoidance task, RT: reaction time; T1: Time 1; T2: Time 2; N: number of valid participants for each task, Time1/Time2. Differences in N reflect differences in sample size at each testing wave and accounting for data cleaning (see Analyses). ¹ Split-half coefficients use Spearman-Brown corrections ² test-retest correlations partial out individual inter-session intervals.

clinical levels of anxiety were associated with an attentional bias in children whilst both clinically anxious adults and those who self-reported elevated anxiety demonstrated an attentional bias for threat (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). Age-related effects may also in part account for the weak associations observed. However, few child studies were included in the meta-analysis and those included were

characterised by mixed age ranges, limiting the assessment of possible age effects on attentional bias. Unfortunately, the age-range tested in the present study was not sufficiently broad to permit investigation of age-related effects. Previous work has investigated whether age-related differences in the development of cognitive inhibitory skills may moderate the emergence of anxiety-linked biases in childhood (e.g., Cognitive inhibition hypothesis; Morren

Table 3
Correlations between anxiety and bias indices from information-processing paradigms.

	1	2a	2b	3a	3b	4a	4b	4c	4d	5a	5b	6a	6b	7	8	9	10
1 Anxiety	-	-.06	-.23	-.04	-.02	.04	.20	-.06	.06	.21	-.25	-.06	-.02	-.13	-.02	-.08	-.06
2a DPT; Angry bias	.16	-	-.16	.08	.06	.11	-.14	.18	-.22	-.06	-.04	.13	.23	.09	-.02	-.04	.06
2b DPT; Happy bias	-.02	.03	-	.06	.10	.17	.10	.07	.12	.07	.00	.00	-.10	.09	-.21	.11	-.02
3a MPT; Angry bias	-.04	-.10	-.10	-	.17	.27	-.10	.04	.09	.02	.18	-.05	-.12	.09	-.13	-.16	-.10
3b MPT; Happy bias	.00	.08	-.13	.23	-	.17	.08	.08	.06	-.03	-.20	-.11	-.13	.15	.05	-.11	-.16
4a VS; Angry bias (absent)	.21	-.05	.06	-.08	-.08	-	-.07	.54*	.11	.18	.11	-.08	.03	.09	-.03	.01	-.04
4b VS; Angry bias (present)	.17	-.03	.11	.11	.09	.32	-	-.03	.35	.06	.06	-.25	-.13	.03	.00	-.06	-.10
4c VS; Happy bias (absent)	.09	.12	-.13	-.03	.08	.61*	.23	-	.17	.14	.02	-.09	.04	.18	.00	.08	-.02
4d VS; Happy bias (present)	.08	.00	-.06	.07	-.02	.63*	.54*	.20	-	.07	-.06	-.13	-.11	-.20	-.08	-.17	-.04
5a Stroop; Angry bias	.04	.12	-.13	-.07	.15	-.03	-.22	-.06	.04	-	.56*	.01	.14	.04	-.21	.00	.13
5b Stroop; Happy bias	.18	.12	-.10	-.04	.19	-.10	-.03	-.06	.04	.51*	-	.06	.13	-.08	.06	.03	-.01
6a Garner; Angry bias	-.01	-.12	-.14	.19	-.01	-.02	.18	.14	.02	-.02	.16	-	.33	.22	-.16	.19	.01
6b Garner; Happy bias	-.03	-.15	-.07	.18	-.03	.01	-.47	.14	-.13	-.04	.08	.61*	-	.09	-.17	.05	.04
7 Morph	-.17	.07	.13	-.13	-.04	.13	-.15	.18	-.05	.03	-.09	-.16	.02	-	.11	.00	.15
8 AAT; Angry bias	-.08	.03	.09	.17	.05	.01	.19	-.06	.05	-.14	-.18	-.11	-.05	-.01	-	.05	.02
9 AAT; Happy bias	-.12	-.05	-.07	.15	.05	-.30	.06	-.05	.15	.11	-.02	-.10	-.06	.02	.18	-	.11
10 AAT; Neutral bias	-.26	-.05	.08	-.06	-.02	-.05	.02	-.11	-.22	-.17	-.08	-.02	-.08	-.16	.25	.10	-

DPT: dot-probe task, MPT: missile-probe task, VS: visual search, AAT: Approach–Avoidance Task.

Note: correlations below the diagonal are between the first measurements of each variable. Correlations above the diagonal are the second measurements. * $p < .0016$ (Bonferroni correction $\alpha = .05/32$). Light grey shading indicates associations between bias scores and anxiety whilst dark grey shading highlights associations between bias scores from the same and related tasks.

et al., 2003) but the evidence thus far has been mixed (Hadwin et al., 2009). Alternatively, young children may be less able to accurately report their emotional symptoms making it harder to detect valid anxiety-related effects in younger samples. To this end, parent- and children-reports frequently show poor concordance (De Los Reyes et al., 2011); concordance between children's and parent's anxiety rating on the RCADS-25 is typically moderate ($r \sim .30$; Muris et al., 2002). As a result, parent-reported anxiety in young children may show stronger associations with information-processing biases. Unfortunately, we did not measure parent-reported anxiety in the current study but this should be the focus of future research.

4.3. Convergence between measures

Bias scores from the range of tasks in the current study showed little convergence with one another. This is in line with several other studies which have found poor convergence between bias indices from similar tasks (Broeren et al., 2011; Dalglish et al., 2003). A lack of convergence could indicate that the tasks measure distinct aspects of information processing (Watts & Weems, 2006). Indeed, the tasks included in the current study were designed to measure varying cognitive processes which might be expected to operate independently (e.g., selective attention using a dot-probe task compared to emotional recognition using a morph task). However, theoretical frameworks in anxiety propose that these are related constructs and so biases measured using different paradigms should be linked (Daleiden & Vasey, 1997; In-Albon & Schneider, 2010). Interestingly, even tasks putatively measuring the same construct like the missile- and dot-probe tasks (both

proposing to measure selective attention) and emotional Stroop and Garner tasks (both designed to measure inhibitory control) showed near-zero convergence with one another, suggesting that these task variants do not all successfully measure the intended pattern of processing selectivity.

5. Implications

The results of the current study raise a number of questions that warrant further investigation. First, although paradigms included in the current study were selected to closely mirror those often used with children, it is unclear how generalisable the current findings would be to studies where the task parameters are modified. It will be important for future research to replicate the results seen here in other sample groups (e.g., with different ages or clinically anxious children) using identical task variants but also alternative task variants (e.g., card-based variants of the Stroop). In the very least, the psychometric properties of specific task variants need to be rigorously examined in individual studies and reported together with the results obtained to aid interpretation of their findings.

Another possible avenue for future research is to try to capture the sources of unreliability when using RT based paradigms with children. One possible contributor to poor reliability in children could be that behavioural responses to emotional stimuli are relatively distal from the actual information-processing mechanisms. Interfering cognitive processes (e.g., distractions) could create both systematic and unsystematic measurement error, especially in children where regulatory skills are less developed and more likely to vary across individuals (Klenberg et al., 2001). It was not possible

to examine age effects in the current study owing to the narrow age range. However, limited past research suggests that RT variability may decrease with age (Broeren et al., 2011). Future research would benefit from formally examining age-related change in children's reaction time variability on a wide range of information-processing tasks in order to establish whether differences in regulatory skills contribute to the poor reliability demonstrated in the current study.

Alternatively, combining RT based tasks with methodologies such as eye-tracking or neurophysiological indices that do not require participants to remember and perform the types of additional responses often required in conventional RT tasks may serve to improve reliability. Eye-tracking represents one such option, which may enable 'online' measurement of attentional deployment during information-processing tasks and thus may reveal more reliable individual differences in attentional responses to emotional stimuli. Studies have identified anxiety-related biases in both initial gaze directions and saccades when both children and adults are presented with emotional stimuli (see In-Albon & Schneider, 2010 for a review). Additionally, psychometric analyses of eye-tracking measures of attentional bias reveal substantial internal consistency ($<.80$) and retest reliability (.43–.79). Electrophysiological indices of activity in brain regions involved in emotional processing also show promise. Some studies have shown that clinically-anxious children and adults show greater amygdala activation than do non-anxious individuals in response to emotional faces (Stein, Simmons, Feinstein, & Paulus, 2007), and that high anxious relative to low anxious individuals show enhanced event-related potentials in response to angry faces on a spatial-cueing paradigm (Fox, Derakshan, & Shoker, 2008). There is also limited evidence for adequate psychometric properties of these indices (Tomarken, Davidson, Wheeler, & Kinney, 1992). Studies using physiological indices as measures of anxiety-related processing biases are in their infancy, and tend to have small sample sizes. However, they present promising methodologies and future studies should aim to assess their psychometric properties in child anxiety samples.

6. Limitations

Poor reliability estimates and inconsistent associations between measures were found in spite of having more than adequate power; estimates ranged from 87% to 99% power to detect moderate effect sizes (.3–.5; Field, 2005) with our smallest sample size (102 children at wave 3). Nevertheless, a number of study-specific limitations are worth considering. First, poor task comprehension could have contributed to low reliability. However, this is unlikely in the current study since instructions were read aloud by researchers and practice blocks ensured full comprehension prior to task commencement. Very low error rates further support the adequacy of the instructions employed in these studies. Second, all data was collected during school hours, and although testing took place in an unused classroom to reduce disruption, the nature of school environments meant some inevitable distractions. Environmental distraction could have introduced measurement error and attenuated reliability. Third, the age range of the current study was relatively narrow (representing a two year interval in middle childhood) and so age-related effects on task performance could not be examined. Additionally, the use of children's self report and an unselected sample may have limited the ability to detect associations with anxiety. However, the range of anxiety scores and the proportion of children with clinically elevated anxiety was comparable to other normative samples (Muris et al., 2002), suggesting sufficient variability to detect associations with information-processing biases. Nevertheless, future research would benefit from a systematic assessment of the reliability of information-processing tasks,

their convergence and associations with anxiety across development and in both unselected and clinically anxious samples.

7. Conclusions

If replicated, the present finding that tasks yielding processing bias measures lack sound psychometric properties when used to assess children has implications for clinical and experimental practice. The observed poor reliability of these tasks in children who differ in levels of self-reported anxiety suggests they may be poorly suited for measuring anxiety-linked differences in emotional processing biases in child samples. Hence, caution should be taken when interpreting results from studies employing such approaches to assess children. It would be useful for future studies investigating anxiety-linked processing biases in children to rigorously examine the psychometric properties of the adopted assessment tasks, and to report these properties together with the findings obtained, to aid interpretations of the findings. While the present findings indicate the potential importance of this approach to child research in this field, the adoption of such good practice would also be appropriate when using such cognitive-experimental tasks to measure anxiety-linked processing biases in adults.

Acknowledgements

This research was funded by an ESRC 1 + 3 PhD studentship awarded to the first author. The corresponding author was funded by the MRC during completion of this research (G0901874 and MR/J011762/1). The authors would like to thank the EDITLab team at the SGDP, especially Gillian O'Neill, Melissa Howard, Lucy Lindley, Michelle Vermeulen, Nina Mikita, Tavia Evans, Rachel Banham, Sneha Paleja, Ayushi Desai and Hiba Kazmi for their assistance with data collection. Special thanks go to Dulwich Hamlet Junior School and Beckenham Junior School for their participation.

References

- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, 133(1), 1–24. <http://dx.doi.org/10.1037/0033-2909.133.1.1>
- Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: automatic and strategic processes. *Behaviour Research and Therapy*, 35(1), 49–58. [http://dx.doi.org/10.1016/S0005-7967\(96\)00069-1](http://dx.doi.org/10.1016/S0005-7967(96)00069-1)
- Broeren, S., Muris, P., Bouwmeester, S., Field, A. P., & Voerman, J. S. (2011). Processing biases for emotional faces in 4- to 12-year-old non-clinical children: an exploratory study of developmental patterns and relationships with social anxiety and behavioral inhibition. *Journal of Experimental Psychopathology*, 2(4), 454–474. <http://dx.doi.org/10.5127/jep.016611>
- Brown, H. M., McAdams, T. A., Lester, K. J., Goodman, R., Clark, D. M., & Eley, T. C. (2013). Attentional threat avoidance and familial risk are independently associated with childhood anxiety disorders. *Journal of Child Psychology and Psychiatry*, 54(6), 678–685. <http://dx.doi.org/10.1111/jcpp.12024>
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25(2), 215–224. <http://dx.doi.org/10.1177/0146167299025002007>
- Daleiden, E. L., & Vasey, M. W. (1997). An information-processing perspective on childhood anxiety. *Clinical Psychology Review*, 17(4), 407–429. [http://dx.doi.org/10.1016/S0272-7358\(97\)00010-X](http://dx.doi.org/10.1016/S0272-7358(97)00010-X)
- Dalgleish, T., Taghavi, R., Neshat-Doost, H., Moradi, A., Canterbury, R., & Yule, W. (2003). Patterns of processing bias for emotional information across clinical disorders: a comparison of attention, memory, and prospective cognition in children and adolescents with depression, generalized anxiety, and posttraumatic stress disorder. *Journal of Clinical Child & Adolescent Psychology*, 32(1), 10–21. <http://dx.doi.org/10.1207/s15374420360533022>
- De Los Reyes, A., Youngstrom, E. A., Swan, A. J., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2011). Informant discrepancies in clinical reports of youths and interviewers' impressions of the reliability of informants. *Journal of Child and Adolescent Psychopharmacology*, 21(5), 417–424.
- Donnelly, N., Hadwin, J. A., Menneer, T., & Richards, H. J. (2010). The use of visual search paradigms to understand attentional biases in childhood anxiety. In: J. A. Hadwin, & A. P. Field (Eds.), *Information processing biases and anxiety: a developmental perspective*. Chichester: Wiley-Blackwell.

- Eide, P., Kemp, A., Silberstein, R. B., Nathan, P. J., & Stough, C. (2002). Test-retest reliability of the emotional Stroop task: examining the paradox of measurement change. *Journal of Psychology*, 136(5), 514–520. <http://dx.doi.org/10.1080/00223980209605547>
- Field, A. P. (2005). *Discovering statistics using SPSS* (2) London: Sage publications.
- Field, A. P., & Lester, K. J. (2010). Is there room for 'Development' in developmental models of information processing biases to threat in children and adolescents? *Clinical Child and Family Psychology Review*, 13(4), 315–332.
- Fox, E., Derakshan, N., & Shoker, L. (2008). Trait anxiety modulates the electrophysiological indices of rapid spatial orienting towards angry faces. *Neuroreport*, 19(3), 259–263. <http://dx.doi.org/10.1097/WNR.0b013e3282f53d2a>
- Gilboa-Schechtman, E., Ben-Artzi, E., Jeczemien, P., Maro, S., & Hermesh, H. (2004). Depression impairs the ability to ignore the emotional aspects of facial expressions: evidence from the Garner task. *Cognition & Emotion*, 18(209–231) <http://dx.doi.org/10.1080/02699930341000176>
- Gullone, E. (2000). The development of normal fear: a century of research. *Clinical Psychology Review*, 20(4), 429–451. [http://dx.doi.org/10.1016/S0272-7358\(99\)00034-3](http://dx.doi.org/10.1016/S0272-7358(99)00034-3)
- Hadwin, J. A., Donnelly, N., French, C. C., Richards, A., Watts, A., & Daley, D. (2003). The influence of children's self-report trait anxiety and depression on visual search for emotional faces. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44(3), 432–444. <http://dx.doi.org/10.1111/1469-7610.00133>
- Hadwin, J. A., Donnelly, N., Richards, A., French, C. C., & Patel, U. (2009). Childhood anxiety and attention to emotion faces in a modified stroop task. *British Journal of Developmental Psychology*, 27(Pt 2), 487–494. <http://dx.doi.org/10.1348/026151008X315503>
- Hadwin, J. A., & Field, A. P. (2010). *Information processing biases and anxiety: a developmental perspective*. Chichester: Wiley-Blackwell.
- Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., et al. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, 68(11), 982–990. <http://dx.doi.org/10.1016/j.biopsych.2010.07.021>
- Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: the approach-avoidance task. *Behaviour Research and Therapy*, <http://dx.doi.org/10.1016/j.brat.2007.08.010>
- In-Albon, T., & Schneider, S. (2010). Using eye tracking methodology in children with anxiety disorders. In: J. Hadwin, & A. Field (Eds.), *Information processing biases and anxiety: a developmental perspective* (p. 129). Chichester, UK: John Wiley & Sons Ltd. <http://dx.doi.org/10.1002/9780470661468.ch6>
- Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, 115(4), 705–714.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 593–602. <http://dx.doi.org/10.1001/archpsyc.62.6.593>, pii:62/6/593
- Klein, A. M., Becker, E. S., & Rinck, M. (2011a). Approach and avoidance tendencies in spider fearful children: the approach-avoidance task. *Journal of Child and Family Studies*, 20(2), 224–231. <http://dx.doi.org/10.1007/s10826-010-9402-7>
- Klein, A. M., Becker, E. S., & Rinck, M. (2011b). Direct and indirect measures of spider fear predict unique variance in children's fear-related behaviour. *Cognition & Emotion*, 25(7), 1205–1213. <http://dx.doi.org/10.1080/02699931.2010.530454>
- Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, 20(1), 407–428. <http://dx.doi.org/10.1207/S15326942DN2001.6>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D., Hawk, S., & van Knippenberg, A. (2012). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24, 1377–1388. <http://dx.doi.org/10.1080/02699930903485076>
- Lau, J. Y., Burt, M., Leibenluft, E., Pine, D. S., Rijdsdijk, F., Shiffrin, N., et al. (2009). Individual differences in children's facial expression recognition ability: the role of nature and nurture. *Developmental Neuropsychology*, 34(1), 37–51. <http://dx.doi.org/10.1080/87565640802564424>
- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95(1), 15–20. <http://dx.doi.org/10.1037/0021-843X.95.1.15>
- Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, 5(1), 119. <http://dx.doi.org/10.1037/1528-3542.5.1.119>
- Mathews, A., Mogg, K., Kentish, J., & Eysenck, M. (1995). Effect of psychological treatment on cognitive bias in generalized anxiety disorder. *Behaviour Research and Therapy*, 33(3), 293–303. [http://dx.doi.org/10.1016/0005-7967\(94\)E0022-B](http://dx.doi.org/10.1016/0005-7967(94)E0022-B)
- Mattia, J. I., Heimberg, R. G., & Hope, D. A. (1993). The revised Stroop color-naming task in social phobics. *Behaviour Research and Therapy*, 31(3), 305–313. [http://dx.doi.org/10.1016/0005-7967\(93\)90029-T](http://dx.doi.org/10.1016/0005-7967(93)90029-T)
- MorphMan 4.0 (2003). Moscow, Russia: STOIK Imaging.
- Morren, M., Kindt, M., van den Hout, M., & van Kasteren, H. (2003). Anxiety and the processing of threat in children: further examination of the cognitive inhibition hypothesis. *Behaviour Change*, 20(3), 131–142. <http://dx.doi.org/10.1375/behc.20.3.131.24833>
- Muris, P., & Field, A. P. (2008). Distorted cognition and pathological anxiety in children and adolescents. *Cognition & Emotion*, 22(3), 395–421. <http://dx.doi.org/10.1080/02699930701843450>
- Muris, P., Meesters, C., & Schouten, E. (2002). A brief questionnaire of DSM IV defined anxiety and depression symptoms among children. *Clinical Psychology & Psychotherapy*, 9(6), 430–442. <http://dx.doi.org/10.1002/cpp.347>
- Overall, J. E., & Woodward, J. A. (1975). Unreliability of difference scores: a paradox for measurement of change. *Psychological Bulletin*, 82(1), 85–86. <http://dx.doi.org/10.1037/h0076158>
- Perez-Olivas, G., Stevenson, J., & Hadwin, J. A. (2008). Do anxiety-related attentional biases mediate the link between maternal over involvement and separation anxiety in children? *Cognition & Emotion*, 22(3), 509–521.
- Richards, A., French, C. C., Nash, G., Hadwin, J. A., & Donnelly, N. (2007). A comparison of selective attention and facial processing biases in typically developing children who are high and low in self-reported trait anxiety. *Development and Psychopathology*, 19(2), 481–495. <http://dx.doi.org/10.1017/S095457940707023X>
- Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, 38(2), 105–120.
- Roy, A. K., Vasa, R. A., Bruck, M., Mogg, K., Bradley, B. P., Sweeney, M., et al. (2008). Attention bias toward threat in pediatric anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(10), 1189–1196. <http://dx.doi.org/10.1097/CHI.0b013e3181825ace>
- Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality*, 19(7), 595–605. <http://dx.doi.org/10.1002/per.554>
- Scott, A., Pearce, D., & Goldblatt, P. (2001). The sizes and characteristics of the minority ethnic population in Great Britain: latest estimates. *Population Trends*, 105, 5–15. www.ncbi.nlm.nih.gov/pubmed/11599123
- Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, 51(3), 339–350. www.ncbi.nlm.nih.gov/pubmed/11599123
- Stein, M., Simmons, A., Feinstein, J., & Paulus, M. (2007). Increased amygdala and insula activation during emotion processing in anxiety-prone subjects. *American Journal of Psychiatry*, 164(2), 318–327. <http://dx.doi.org/10.1176/appi.ajp.164.2.318>
- Stirling, L. J., Eley, T. C., & Clark, D. M. (2006). Preliminary evidence for an association between social anxiety symptoms and avoidance of negative faces in school-age children. *Journal of Clinical Child and Adolescent Psychology*, 35(3), 431–439. <http://dx.doi.org/10.1207/s15374424jccp3503.9>
- Strauss, G. P., Allen, D. N., Jorgensen, M. L., & Cramer, S. L. (2005). Test-retest reliability of standard and emotional Stroop tasks. *Assessment*, 12(3), 330–337. <http://dx.doi.org/10.1177/1073191105276375>
- Telzer, E. H., Mogg, K., Bradley, B. P., Mai, X., Ernst, M., Pine, D. S., et al. (2008). Relationship between trait anxiety, prefrontal cortex, and attention bias to angry faces in children and adolescents. *Biological Psychology*, 79(2), 216–222. <http://dx.doi.org/10.1016/j.biopsycho.2008.05.004>
- Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Kinney, L. (1992). Psychometric properties of resting anterior EEG asymmetry: temporal stability and internal consistency. *Psychophysiology*, 29(5), 576–592. <http://dx.doi.org/10.1111/j.1469-8986.1992.tb02034.x>
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249. <http://dx.doi.org/10.1016/j.psychres.2008.05.006>
- Waters, A. M., Lipp, O. V., & Spence, S. H. (2004). Attentional bias toward fear-related stimuli: an investigation with nonselected children and adults and children with anxiety disorders. *Journal of Experimental Child Psychology*, 89(4), 320–337. <http://dx.doi.org/10.1016/j.jecp.2004.06.003>
- Waters, A. M., Mogg, K., Bradley, B. P., & Pine, D. S. (2008). Attentional bias for emotional faces in children with generalized anxiety disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(4), 435–442. <http://dx.doi.org/10.1097/CHI.0b013e3181642992>
- Watts, S. E., & Weems, C. F. (2006). Associations among selective attention, memory bias, cognitive errors and symptoms of anxiety in youth. *Journal of Abnormal Child Psychology*, 34, 838–849.